DOCUMENT RESUME

ED 415 622 EC 306 102

AUTHOR	Ysseldyke, Jim; Thurlow, Martha; Erickson, Ron; Haigh, John; Moody, Mark; Trimble, Scott; Insko, Bill
TITLE	Reporting School Performance in the Maryland and Kentucky Accountability Systems: What Scores Mean and How They Are Used. State Assessment Series: Maryland/Kentucky Report 2.
INSTITUTION	Maryland State Dept. of Education, Baltimore.; National Center on Educational Outcomes, Minneapolis, MN.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	1997-09-00
NOTE	32p.
CONTRACT	H023F50008
AVAILABLE FROM	National Center on Educational Outcomes, University of Minnesota, 350 Elliott Hall, 75 East River Road, Minneapolis, MN 55455; telephone: 612-624-8561; fax: 612-624-0879; World Wide Web: http://www.coled.umn. edu/NCEO (document may be copied free of charge, additional print copies $8).
PUB TYPE	Reports - Descriptive (141) -- Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Standards; Accountability; *Disabilities; *Educational Assessment; Educational Testing; Elementary Secondary Education; Evaluation Methods; Outcomes of Education; State Programs; *Student Evaluation; *Student Participation; *Test Use
IDENTIFIERS	Kentucky; Maryland; Testing Accommodations (Disabilities)

ABSTRACT

This report describes how assessments are scored and results are reported in the inclusive assessment and accountability systems in Maryland and Kentucky. Specific topics include: (1) how school performance is reported to the public; (2) what components go into the reported performance of schools; (3) how scores are obtained for students in the regular assessments; (4) how scores are obtained for students who participate in alternate assessment systems; (5) how scores are aggregated from regular and alternate assessment systems; (6) what current issues the assessment programs in both states are facing; and (7) what the plans are for the future. These topics are discussed for each state separately and then the two states are compared. Differences between the two states include the establishment of benchmark performance standards, Maryland's use of test performance in graduation requirements, Maryland's detailed annual report of the performance of each of the local school systems, and Maryland's development of a separate set of academic expectations for students in their alternate assessment programs. Commonalities between the two states are also discussed. An appendix includes a sample of state summary and disaggregated data, and summary data for each school system in Maryland. (CR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 415 622

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- [x] This document has been reproduced as received from the person or organization originating it.
- [ ] Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

EC 306102

STATE ASSESSMENT SERIES
Maryland/Kentucky
Report 2

Reporting School Performance in the Maryland and Kentucky Accountability Systems: What Scores Mean and How They Are Used

The Maryland-Kentucky-NCEO Assessment Project

Prepared By:

Jim Ysseldyke • Martha Thurlow • Ron Erickson
National Center on Educational Outcomes, University of Minnesota

John Haigh • Mark Moody
Maryland Department of Education

Scott Trimble • Bill Insko
Kentucky Department of Education

September 1997

The Maryland-Kentucky-NCEO Assessment Project encompasses a comprehensive array of research studies on assessment accommodations and alternate assessment approaches that facilitate inclusion of all students in statewide assessment programs. The planned research program include:

- comparative studies involving a qualitative examination of the assessment systems in Kentucky and Maryland;
- studies that involve secondary analyses of existing data bases to address critical technical and implementation issues in the assessment of students with disabilities; and
- experimental field studies that involve the collection of new data to address several technical issues crucial to the development and modification of state assessment policies.

This project is supported, in part, by a grant to the Maryland State Department of Education from the U.S. Department of Education, Office of Special Education Programs (Grant #H023F50008). Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $8.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.624.8561 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

Overview

Calculating the scores from an assessment is not always a clear-cut or easily understood process. Large-scale assessments frequently use a variety of complex techniques, such as matrix sampling, in which no student takes the entire test, yet the scores of students are aggregated to estimate average performance on the test. Item Response Theory (IRT) modeling and scaling techniques are used to determine item characteristic curves and difficulty levels. Now, with the move toward high standards, scoring techniques that focus on comparing students with one another no longer are viewed as appropriate; student performance must be compared to established standards. Increasingly, new scoring techniques are used to reflect the extent to which students and schools are meeting desired levels of performance. Assessment systems have gone far beyond the simple use of total raw scores, stanines, grade equivalents, standard scores, and even normal curve equivalents.

Against the backdrop of sophisticated test development and analysis techniques, it is often difficult to understand how scores are calculated in many current statewide assessment programs. When dealing with accountability systems that are complex and that attempt to include all students (including students with disabilities) in the system, it is often even more difficult to understand how scores are obtained so that comprehensible reports can be developed.

Kentucky and Maryland are two states regarded as having the most inclusive systems of educational accountability (Ysseldyke, Thurlow, Erickson, Gabrys, Haigh, Trimble & Gong, 1996). Both states have adopted the approach that assessments are to be taken by all students, including a very small number of students needing an alternate assessment. Both states have school accountability systems in which there are significant consequences for schools based on student test performance and other indicators of success (e.g., increased attendance rates, decreased dropout rates).

Among the frequently raised questions about the scoring of assessments for the accountability systems in these two states are:

- How do these states use assessment results to describe the performance of their schools and districts?
- When using a single index to describe the performance of a school, how are the individual components of that index obtained?
- How is the performance of an individual student scored?
- How are the scores of students in alternate assessment systems (for students

with severe disabilities) combined with the scores for students in the regular assessment system? How do we know that they mean the same thing?

These and related questions are the focus of this paper. In order to clarify scoring and reporting in the Kentucky and Maryland accountability systems, this report will address:

- How school performance is reported to the public.
- What components go into the reported performance of schools.
- How scores are obtained for students in the regular assessments.
- How scores are obtained for students who participate in alternate assessment systems.
- How scores are aggregated from regular and alternate assessment systems.
- What current issues the assessment programs in both states are facing.
- What the plans are for the future.

We first address these topics for each state separately. Then, we offer some comparative remarks about these two states and their educational accountability systems.

Measuring and Reporting School Performance in Kentucky

The Kentucky Education Reform Act (KERA) of 1990 formed the basis for massive change in the state's educational system. This massive reform was enacted by the Kentucky General Assembly as a result of a lawsuit brought by the Coalition for Better Education (CBE), which represented approximately 60 of the state's 176 school districts. The successful 1988 lawsuit found the state's funding mechanisms inequitable and mandated that the educational system be redesigned. The reform called for top-down and bottom-up systemic change in finance, governance, curriculum, and assessment.

KERA established six goals for the schools of the Commonwealth: (1) expect a high level of achievement of all students, (2) develop students' abilities in six cognitive areas, (3) increase school attendance rates, (4) reduce dropout and retention rates, (5) reduce physical and mental health barriers to learning, and (6) increase the proportion of students who make a successful transition to work, postsecondary education, and the military.

Immediately needed under the requirements of the Act was an assessment system capable of measuring progress toward the goals, primarily the academic expectations reflected in the first two goals. Through a competitive process, the Kentucky Department of Education selected Advanced Systems in Measurement and Evaluation as the contractor for the assessment program, which came to be known as the Kentucky Instructional Results Information System (KIRIS).

The contents of the KIRIS assessment components were influenced primarily by the direction of content area advisory committees, with members drawn mostly from classrooms, schools, professional education organizations, higher education, community groups, and the Kentucky Department of Education. The KIRIS assessment, which has been administered annually since the spring of 1992, has included three types of assessment tasks:

Assessment tasks involving portfolios. Each student in grades 4, 8, and 12 is required to assemble a Writing Portfolio and a Mathematics Portfolio (as of the 1994-95 school year Mathematics Portfolios are required in grade 5, rather than grade 4). These portfolios represent collections of the student's best work developed over time in conjunction with support from teachers, peers, and parents. The portfolios are scored by local teachers, and the scores are reported to the Kentucky Department of Education for use in the accountability assessment. Mathematics portfolios will not be included in the baseline calculation for 1996-97 and 1997-98, but will be included for instructional purposes in 1997-98, and for accountability purposes in 1998-99.

Assessment tasks involving performance events. Students participate in assessment tasks that require them to use knowledge and skills learned in school to produce a product or solve a problem. Rather than recall facts, students apply what they have learned to a real (or real-life simulated) situation. Performance event tasks, which involve both group and individual work, are based on manipulatives or other materials and take about an hour each for completion. Performance event tasks are administered by test administrators hired by Advanced Systems in Measurement and Evaluation. For 1996-97 and beyond, performance events enter a research and development phase because of technical considerations. Until this is complete, they will not be included in the accountability index.

Assessment tasks involving open-ended questions. Students respond to open-ended questions requiring extended written responses. The focus is on higher-order thinking skills, solving multi-step problems, and using reasoning, analytical, and written communication skills.

Assessment tasks involving machine-scorable questions. In 1992-94 students also answered a section of multiple choice questions, although these were not used for accountability purposes. Beginning in 1994-95, KIRIS included a section of other item types being evaluated for possible inclusion in the future. Beginning in 1996-97, a section of multiple choice questions will be included in each content area, and phased in for accountability purposes.

KIRIS also monitors school progress in terms of non-cognitive indicators such as school attendance rates, dropout and retention rates, and the proportion of students who make a successful transition to work, postsecondary education, or the military.

KIRIS assessments are administered every year to students in grades 4/5, 7/8, and 11/12. (Beginning in the 1996-97 school year, administrations for certain subjects at grades 5 and 7 were implemented to reduce the total amount of testing time for students, and still obtain the range of information needed for the Kentucky accountability system.) Information gathered through the KIRIS assessment system forms the basis for deriving a single statistic describing a school's performance, called the *school accountability index*. This single index is derived from assimilating measures in both cognitive and noncognitive areas of student performance; however, the cognitive component remains the major component in calculating each school's index.

As reported elsewhere (Ysseldyke et al., 1996), each school in Kentucky is working toward an "improvement goal." If the school accountability index is more than 1 point above that goal, the school receives a financial reward, to be spent in any way agreed upon by the majority of its certified staff. If the school is above the improvement goal, but by less than 1 point, the school is not eligible for a reward. If the school accountability index is below the goal, the school receives assistance. The first time a school does not perform at its goal, it must develop a school improvement plan, and receives some funds to support these improvements. The second time a school does not perform at its goal, it is designated a "school in decline" and receives the services of a master (distinguished) teacher. If the school fails to reach its improvement goal a third time, it is declared a "school in crisis." At this point, a school may be taken over by the state, or its administrative personnel replaced, or the entire school reconstituted. While master teachers have been available for schools not showing sufficient improvement, no serious consequences (e.g., firing administrators, replacing teachers) have yet been imposed. A school can go into "decline" by scoring below past performance (but not by 5 or more points). It is also possible for a school to immediately be declared a "school in crisis" (if its score is 5 points or more below its improvement goal).

At the state level, Kentucky does not have a student-level accountability system. That is, there is no state-level assessment that has the purpose of, by itself or with other information, determining whether a student will be promoted from one grade to the next, or whether a student will graduate from high school. Nevertheless, within local school systems there has been interest in being able to determine student-level scores for various purposes.

How School Scores Are Reported

School scores are the scores of primary interest in Kentucky. However, it is not absolute or average scores that are of interest, but rather, the percentage of students in the school meeting different levels of proficiency across subject areas, compared to the school-specific progress. Student work within Kentucky's assessment program is scored and assigned to one of four different levels, or standards of performance: *novice, apprentice, proficient,* or *distinguished.* These standards are set specific to each content area and grade level. Table 1 offers definitions of these four levels of student performance.

Table 1. Levels of Student Performance in Kentucky Instructional Information System (KIRIS)

Level of Performance	Definition
Distinguished	A level of performance above proficient for "that small percentage of students who exceed even the Proficient standard."
Proficient	The desired level of performance, which "will allow the student to be competitive in the economic and social environment of the next century."
Apprentice	A level of performance that is "intermediate between Novice and Proficient."
Novice	A level of performance that demonstrates few or none of the qualities of proficiency.

Note. Definitions are from Trimble, 1994, p. 47.

Baseline scores were first obtained for Kentucky's schools in 1991-92. (During baseline, approximately 10 percent of all Kentucky students performed at the proficient level in reading, math, science, and social studies.) The accountability index calculated for the baseline year was used to determine the improvement goal, which reflected the growth required for performance to be considered adequate. The improvement goal was obtained by subtracting the baseline level from 100 (an index level that could be attained if all students were proficient in all areas, and noncognitive data were perfect), divided by ten (accountability cycles are four-year periods).

For each cycle, a new school accountability index is obtained and used to plot a revised goal for the school. The average of the first two years of a cycle is often referred to as a baseline index; the average of the second two as a growth index.

School accountability scores are based on the performance of *all* students in a school. In Accountability Cycle 1 (1991-92/1993-94) , school scores did not include the performance of students with severe disabilities because the alternate assessment system for these students was still under development. They did include the scores of students with disabilities who participated in the regular assessments, either with or without accommodations. In the second cycle (1992-93/1995-96) students with severe disabilities were included. If students were served in special schools, scores from these students were assigned to the students' home schools, regardless of where the students were receiving instruction (e.g., residential placement). Regular schools hosting special programs could keep all scores so long as data were handled consistently over time.

Components of Reported School Scores

During Kentucky's first accountability cycle (i.e., 1991-92 to 1993-94) students participating in the regular assessment system in Kentucky were involved in three kinds of assessments: a transitional assessment, performance events, and writing portfolios. Although multiple choice items were available in the transitional assessment, only the open-ended items were actually used in the accountability system. By the 1993-94 school year, the transitional assessment included five common open-ended items and 24 matrix sampled items (2 in each of 12 different forms), totaling 29 items per content area (i.e., reading, mathematics, science, and social studies) for each grade tested.

The original intent in changing the KIRIS assessment over time was to increase the use of open response items and performance events, designed to better reflect the goals of reforms in instruction and learning environments. They emphasize applying skills to produce products and require that groups of students work together (Trimble & Forsaith, 1995). By 1993-94, the performance events were becoming more interdisciplinary, frequently crossing mathematics and science, or social studies, and adding in arts and humanities.

Portfolios for KIRIS have been implemented in writing since 1991-92 and in mathematics since 1993-94. (Mathematics portfolios have been put on hold for accountability purposes during the 1996-97 and 1997-98 testing cycle.) While the portfolios are to represent "best work" of the student, they are standardized in the sense that each portfolio is to contain specific types of entries, with a set of standards identified a priori for judging the portfolios.

Alternate portfolios are the only form of assessment for students with moderate to severe disabilities, students who "are not pursuing a high school diploma," or otherwise not participating in the general curriculum (Trimble & Forsaith, 1995, p. 631). Yet, the information for these students contributes to the aggregate scores in the same way as do the scores from other students.

In addition to student performance, the noncognitive indicators contribute approximately 16 percent to school scores (Trimble & Forsaith, 1995). The definitions of these indicators result in their values being relatively consistent. It has been argued that this consistency results in a real impact that is less than 16 percent (Trimble & Forsaith, 1995).

Scoring Student Performance

Standards of performance were established by reaching consensus on what needed to be accomplished. *Proficient* performance was the desired standard. Three other levels of performance (*novice*, *apprentice*, and *distinguished*) were defined relative to this standard (refer to Table 1). Proficient is described as "the desired level of performance, that which will allow the student to be competitive in the economic and social environment of the next century" (Trimble, 1994, p. 47).

The decision about what performance level a student is demonstrating depends on the nature of the assessment item. Most standards are dependent on teacher judgment. These are based on extensive training and follow-up checks of reliability. For open-response items, both raw-score distributions and observed characteristics of the distributions resulting from Item Response Theory (IRT) analyses were used to determine score alignment with performance level. Students who do not take a component of the assessment are assigned to the "novice" level for that component, and this is entered into the accountability system.

The values assigned to the four performance levels were transformed to ones in which a score of 1.00 reflected the proficient level, the target level for KIRIS. Using this transformation, the novice, apprentice, proficient, and distinguished levels are represented by the values 0, 0.4, 1.0, and 1.4, respectively. These are the weights used in calculating each school's accountability index.

For students with more severe cognitive disabilities, the Alternate Portfolio was developed to reflect the same set of learner outcomes as for all other students. As Kleinert, Kearns, and Kennedy (1996) noted:

> The expectation of *accessing information* for a student with severe disabilities may be demonstrated by that student's skills in appropriately requesting needed assistance across multiple school and community settings. The expectation of

> *using technology effectively* could be evidenced through appropriate assistive technology applications (e.g., using an augmentative communication device, or operating a computer program through single switch access). (p. 7)

To determine scoring standards for the Alternate Portfolio, the Alternate Portfolio Advisory Committee met with people throughout the state to identify sample portfolios (from those developed in 1992-93) that could be used as benchmarks of the four performance levels. These were selected on the basis of what was considered to be "best practices." Six scoring standards were used in this delineation (see Table 2). A clustering of the six standards was used to assign single scores of novice, apprentice, proficient, or distinguished to students' portfolios.

The score of a student on the Alternate Portfolio contributes to a school accountability index in a way equivalent to that of a student in the regular assessment. The underlying philosophy is that the impact of a student's performance within the alternate portfolio process must have the same impact on the final index as the student in the general population. This is accomplished by assuming that the alternate portfolio is an indication of an instructional program's effectiveness across *all of the various content areas* addressed in the regular assessment program. Therefore, before percentages are calculated for each of the four performance levels in any particular content area (e.g., reading, mathematics, science) the performance distribution of students engaged in the alternate portfolio is added to the count used for determining those percentages.

Aggregating Performance Scores

The aggregation of most interest in Kentucky is the school accountability index which, in Accountability Cycle 1, was the average of the cognitive and noncognitive indicators. The *baseline score* for a school reflects the percentage of successful students. The *improvement goal* reflects what the percentage must be over the next two years. A single score for a school is calculated by combining cognitive and noncognitive indicators included in the accountability index. Cognitive indicators include student performance on reading, mathematics, science, social studies, and writing. Noncognitive indicators include attendance rate, retention rate, dropout rate, and transition success.

The manner in which the indicators are combined to form a school score seems somewhat complex on first view, but really is a logical approach to determining ways to combine variables whose relevance and influence vary in different content areas and at different grade levels. Overall, however, the guiding rule in combining indicators for the accountability index is that student success in reading, math, science, social studies, and writing is combined with one index of the noncognitive indicators. The contributions of the noncognitive indicators to the noncognitive index vary by grade. This is because the importance of the different indicators is assumed to vary at different grades. For example, at grade 4, neither the dropout rate nor the

Table 2: Scoring Standards for Alternate Portfolio System

Standard 1	Documented performance of Kentucky's learner outcomes identified for all students (e.g., ability to communicate effectively, to use quantitative or numerical concepts in real life problems, to use effective interpersonal skills, etc.), evidence across the major life domains (personal management, recreation/leisure, and vocational) that are the focus of a community-referenced curriculum.
Standard 2	The student's ability to plan, initiate, monitor, and evaluate his/her own performance within and across entries.
Standard 3	The use of appropriate technology and adaptive/assistive devices within age-appropriate, functional activities, systematic evidence of student choice-making throughout the day, as well as a broad application of different entry types (e.g., investigations and discoveries, projects, instructional programs).
Standard 4	Student outcomes evidenced across multiple school and community settings. For elementary-age students, priority is given to performance in a wide variety of integrated or inclusive school settings. For older students, community-based performance is given increasing emphasis in conjunction with integrated school and class settings.
Standard 5	The degree of student independence in performance, and for those students with more severe disabilities, assistance provided via natural supports, such as peer buddies, peer tutors, and co-workers in job sites, as opposed to assistance provided by paid staff only.
Standard 6	The development of peer interaction skills and mutual friendships with typical peers. [One of the most important, and yet difficult-to-measure, dimensions in the Alternate Portfolio is exactly what constitutes clear evidence of mutual friendships with typical peers.]

Note. Information in this table is taken directly from Kleinert, Kearns, and Kennedy (1996).

transition to adult life success rate is considered relevant; thus, only attendance rate (considered more important and assigned an 80% contribution) and retention rate (assigned a 20% contribution) are included. At grade 8, attendance rate (40% contribution), retention rate (40% contribution), and dropout rate (20% contribution) are all considered relevant, while transition success is not. At grade 12, all indicators are relevant, but not equally so, with attendance rate assigned a 20% contribution, retention rate a 5% contribution, dropout rate a 37.5% contribution, and transition to adult life success a 37.5% contribution. Student success in each content area is

derived by applying an IRT model to all scores, to bring them to the same scale. This entire scale forms the basis for holistic decisions about a student's performance being classified as either novice, apprentice, proficient, or distinguished.

The accountability index for a school recognizes and includes students at different performance levels through the weights assigned by the Board and transformed for the index. In other words, it does not simply count the students who are proficient and above, but rather adds in the relative contribution of those at the apprentice and distinguished levels as well. Thus, the accountability index for an elementary school reflects performance on reading + math + science + social studies + writing + a noncognitive index, by multiplying the percent of students at each performance level times the weight of that level, for each of the components of the indicator, adjusted for a contribution level. Students in the alternate assessment are counted in the same way, with scores weighted in the same way as other students, when the accountability index is derived. In Table 3, a hypothetical school's example displays what would be a typical spread of students in different content areas and on different types of items. The calculations show in a simplified form the way in which the school's accountability index is calculated.

Table 3. Example of Accountability Index Calculation for School A (Grade 4)

Area	Novice (x 0)	Apprentice (x .4)	Proficient (x 1.0)	Distinguised (x 1.4)
Reading	35%	45%	15%	5%
Math	30%	30%	30%	10%
Science	45%	30%	20%	5%
Social Studies	30%	30%	30%	10%
Writing	40%	40%	15%	5%

Noncognitive
Attendance 95%
Retention 1% [1.00 - .01 = .99]

Reading Index = (0 x .35) + (.4 x .45) + (1.0 x .15) + (1.4 x .05) = 0 + .18 + .15 + .07 = .40 x 100 = 40.0
Math Index = (0 x .30) + (.4 x .30) + (1.0 x .30) + (1.4 x .10)] = 0 + .12 + .30 + .14 = .56 x 100 = 56.0
Science Index = (0 x .45) + (.4 x .30) + (1.0 x .20) + (1.4 x .05)] = 0 + .12 + .20 + .07 = .39 x 100 = 39.0
Soc St Index = (0 x .30) + (.4 x .30) + (1.0 x .30) + (1.4 x .10)] = 0 + .12 + .30 + .14 = .56 x 100 = 56.0
Writing Index = (0 x .40) + (.4 x .40) + (1.0 x .15) + (1.4 x .05)] = 0 + .16 + .15 + .07 = .38 x 100 = 38.0
Noncognitive Index = .8 [.95] + .2 [.99] = .760 + .198 = .958 x 100 = 95.8

School Accountability Index = [.40 + .56 + .39 + .56 + .38 + .958] / 5 = 3.248 / 5 = .650 x 100 = 65.0

Current Issues and Future Plans

Kentucky has been very open to study by various research and policy entities. Among those studying Kentucky's assessment system as a whole, including scoring, is the Joint Center for Education Policy (at the University of Kentucky and the University of Louisville). In addition, the Office of Educational Accountability, an office of the state legislature, continues to monitor the effects of the assessment and overall accountability system.

The Kentucky assessment division has been responsive to input from various research and policy entities, with the result being a system that has changed and evolved over time. Among the changes that have occurred are the removal of scores from performance events as they are studied; the addition of a norm-referenced test in the 1996-97 school year (although scores were not to be included in the accountability system); and the inclusion of both writing portfolios and on-demand writing prompts in the accountability equation for the 1996-97 school year. Before the current year, only writing portfolios had been included in the calculation of a school's accountability index. In contrast, mathematics portfolios are not being included in Cycle 3 accountability scores, but are expected to be reintroduced in Cycle 5.

Within a context of considerable change occurring to this state's assessment system, it is not unexpected that the scoring system would need to be adjusted as well. Throughout all this, Kentucky has been vigilant in assessing the reliability of scores assigned to student performance and in the implementation of auditing procedures for ensuring that students are participating in the appropriate assessments.

Education agency officials in Kentucky continue to refine and retool the various components of their state's sweeping reform effort. Particular attention is currently being paid to the issue of supporting teachers and schools in aligning their local curriculum to the expectations reflected in the state assessment program. As noted earlier, administration of a norm-referenced test will be implemented. During Cycle 4, additional multiple choice items (both matrix and common) are being added to the accountability assessments. In addition to this, testing officials have split the testing requirements across pairs of grade levels (e.g., 4th and 5th grade; 7th and 8th grade) in order to reduce the testing burden on students.

Measuring and Reporting School Performance in Maryland

Maryland's focus on school performance and standards began in the late 1980s when the Governor's Commission on School Performance reported that the state lacked an accountability system that could produce good information on how students in Maryland were doing and who should be accountable for changing any evident poor student performance. In 1990 the Maryland School Performance Program (MSPP) was established by the Maryland State Board of Education as the vehicle to move toward a high quality educational system for *all* of Maryland's students. During that year, representatives of numerous groups from across the state (e.g., teachers, parents, administrators) worked to reach consensus on performance areas for which schools should be held accountable.

Student academic performance in Maryland is measured through two assessment programs that are unique to the state. The Maryland School Performance Assessment Program (MSPAP) measures higher-order thinking processes and the application of knowledge and skills to real-world situations as a tool for school improvement and an overall measure of students' knowledge accumulated over several years of schooling. The MSPAP is a single, performance-based test covering mathematics, reading, writing, science, language usage, and social studies. Students in grades 3, 5, and 8 are randomly assigned to one of three clusters per school grade in May of each year. These clusters are composed of portions of the entire MSPAP instrument; consequently, a complete MSPAP score does not exist for any individual student. The assessment takes approximately nine hours of engaged testing time over five days, and includes open-ended questions, essays, and performance events based on Maryland's Learner Outcomes.

An estimated 350 teachers worked with Maryland State Department of Education personnel and a test publisher, CTB Macmillian/McGraw Hill, to develop specifications and the scoring rubrics for these assessments; an additional 600 teachers were hired and trained by another test contractor, Measurement Incorporated, to score them over the summer. The results of the assessments produced scale scores in reading, mathematics, writing, science, social studies, and language usage. These scale scores align with five levels of proficiency. Each proficiency level describes what a student at the level is able to do. Student performance determined to be at Proficiency Level 3 is recognized by the State Board as "satisfactory," and performance assessed as being at Proficiency Level 2 or better is considered "excellent." These proficiency levels have been established and are refined on an annual basis for the MSPAP assessment (see Atash, 1994). Table 4 displays the ranges of scale scores that fall into the various proficiency levels for third graders on the 1994 MSPAP.

Results from the MSPAP testing program are only used in holding schools accountable, while a second assessment program also provides for student-based, as well as school-based, accountability. The Maryland Functional Testing Program (MFTP) includes four basic

Table 4. Grade 3 Proficiency Levels and Corresponding Scale Score Ranges on the 1994 MSPAP

Scale Score Ranges by Subject						
Proficiency Level	**Reading**	**Writing**	**Language Usage**	**Mathematics**	**Science**	**Social Studies**
1	620-700	614-700	620-700	626-700	619-700	622-700
2	580-619	577-613	576-619	583-625	580-618	580-621
3	530-579	528-576	521-575	531-582	527-579	525-579
4	490-529	350-527	350-520	489-530	488-526	495-524
5	350-489	*	*	350-488	350-487	350-494

* Indicates proficiency levels for which cut scores could not be established. These cut scores will be established on future editions of MSPAP.

competency tests that students must pass to receive a Maryland high school diploma. Three of the tests (reading, mathematics, and citizenship) are multiple choice tests; the fourth test, the Maryland Writing Test, is a holistically-scored, direct writing assessment. Although the functional tests have no time limits, the reading, mathematics, and citizenship tests take approximately one hour of engaged testing time. The writing test requires a total of approximately two to three hours over a two-day period. Computer adaptive versions of the reading and mathematics tests take approximately 20 to 30 minutes. The tests formerly were given for the first time in ninth grade; new graduation requirements now permit them to be given as early as grade six. The tests are scored on a Pass or Fail basis, and schools are evaluated on the basis of the percentage of students passing these tests at the end of the ninth and eleventh grades.

Efforts are currently underway within Maryland to eventually retire the MFTP testing program, and replace it with a High School Assessment (HSA) Program, a battery of 10 end-of-course examinations intended to assess secondary students' mastery of core learning goals in English, mathematics, science, and social studies. Although these tests are in the initial stages of development, it is anticipated that the instruments will include short answer, multiple choice, and essay-type items, and that a student's performance on these exams will be linked to a Maryland high school diploma. The first "no fault" administration of these assessments is scheduled for January 1999, in preparation for putting the requirement in place for 9th grade students in the 2000-01 school year.

Maryland recognized that a modified assessment system was needed for a relatively small number of students with more challenging disabilities, and gave special consideration to students with

disabilities fitting one of two profiles. The first type of student participates in the regular MSPAP, but completes the performance assessment tasks with accommodations developed by a team of special education teachers working in collaboration with general education personnel. The second type of student participates in a curriculum focused on functional life skills and is exempted from taking the MSPAP. This group of students is estimated to be less than 5% of all students with disabilities. A state advisory committee (comprised of special education teachers, representatives of advocacy organizations, state education agency staff, and researchers from the University of Maryland) developed a comparative set of outcomes that are appropriate for these students. Currently, this team is developing various assessment and reporting systems to ensure comparable accountability with the regular assessment program. The project, called the Independence Mastery Assessment Program (IMAP), is currently underway in nine of Maryland's local school systems, with projections of increasing this number to 15 districts during the 1997-98 school year, and achieving statewide implementation by the year 2000.

Components of Reported School Performance

A central component of the MSPP is the collection and reporting of information from each school and district in the state. This information is classified as either (1) student performance data or (2) supporting information. The student performance measures currently include assessed student knowledge on the Maryland Functional Testing Program (MFTP) and the Maryland School Performance Assessment Program (MSPAP), along with student participation indices (i.e., attendance rates and dropout rates). Supporting information includes variables related to student population characteristics (enrollment and student mobility); kindergarten completion; numbers of students receiving special services; high school program completion; Grade 12 documented decisions; and other factors such as financial information, staffing, instructional time, and the results of a norm referenced assessment (the Comprehensive Test of Basic Skills/5, given to a sample of students in grades 2, 4, and 6 in each local school system). This supporting information is intended to provide the context for judging each school's growth from year to year.

How School Performance Data are Reported and Used

An emphasis on the public reporting of results is central to the MSPP and its approach to educational accountability. The primary focus of accountability within the state remains on the school; the school district and the state are viewed as support systems to the efforts of schools. On an annual basis, student performance and supporting information is reported at each level of accountability: school, school system, and the state. The *Maryland School Performance Report, State and School Systems,* is published by the Maryland State Department of Education and

includes state summary and disaggregated data and summary data for each school system in the state (see Appendix for an example of this summary). A similar report is published by each local school system, and includes summary and disaggregated data for the local system and for each school in that system. Disaggregated data are reported by gender and race/ethnicity for all student performance data-based areas. School districts and individual schools may add data-based areas and standards that reflect local interests. Actual examples include the extent of advanced placement testing, the number of parent conferences, and the number of volunteers per school.

The data within these reports provide a valuable profile for local committees called School Improvement Teams (SITs). These teams are required of all Maryland schools and are charged with guiding the development of a School Improvement Plan (SIP), to ensure that all students have an opportunity to achieve the outcomes established by the state. The proper use of these data can guide and improve a school's instructional and organizational activities.

The State Department of Education monitors progress of each school annually under an accountability policy know as *reconstitution*. This provision requires that a school not meeting standards must eventually exhibit progress toward those standards. Specific regulations state that a school may be eligible for reconstitution if: (1) it does not meet all the standards and is "below satisfactory and declining" in meeting the appropriate standards, or (2) it does not meet all the standards and is not making "substantial and sustained" improvement through the implementation of a school improvement plan. The specific standards used as the basis for determining a school's reconstitution eligibility differ by grade levels served (see Table 5). As of February 1997, seven high schools, nine middle schools, and 36 elementary schools had been identified as eligible for reconstitution.

Table 5. Specific Standards Monitored for Determining Reconstitution Eligibility

Type of School	Standards Used as Basis for Possible Reconstitution
Elementary Schools	Attendance Rate MSPAP Performance for Grades 3 and 5
Middle Schools	Attendance Rate MSPAP Performance for Grade 8 MFTP Results (Taken in high school and reflected back to appropriate middle school)
High Schools	Attendance Rate Dropout Rate MFTP Results

The process of reconstitution is conducted in several steps, and involves local school system authorities and state education officials. By January 15th of each year, the State Superintendent notifies each local school system of those schools that are failing to meet state standards and are either declining or not making adequate progress. The local school system must then provide the state education agency with a reconstitution proposal, outlining a plan to address the school's areas of need. If approved by the State Board of Education, a more specific transition plan with specific activities and deadlines is required of the local school system by May 15th. A longer term, reconstitution plan is required by January 15th of the following year.

No growth, or continued movement in a downward direction, ultimately could lead to the replacement of a school's administration, staff, or instructional program. However, low performing schools first get technical assistance and additional funding to assist in improving their performance. In fact, the final intent is to allow schools that are successful to help those that are not. As a step in facilitating this improvement strategy, a recognition program for high performing schools was implemented by the Maryland General Assembly in November 1996. The program provides monetary rewards for schools at different levels of performance that improve their performance over two or more years, and provides public recognition for schools that improve over one year.

Setting Standards for School Performance

A critical step in the development of the MSPP was the process of setting standards for the student performance areas against which all schools are measured. Standards have been established for each of the student performance indicators included in the annual reports. These standards define *satisfactory* and *excellent* levels of performance for schools, school systems, and the state to reach by the year 2000. Setting standards for student performance was completed in three phases: (1) the recommendation of a performance range for each measure by a Standards Committee (consisting of 17 members, including representation from 11 local school systems and the Maryland State Department of Education); (2) the modification or approval of that range by a Standards Council (consisting of 12 members, including representation of local education agencies, local boards of education, the state teachers' union — along with a large local union), business interests, students, and the state legislature; and (3) the final adoption of standards by the State Board of Education, following public review and comment. Standards related to pass rates on the Maryland Functional Tests, average daily attendance, and dropout rate were adopted in August 1990. Based on two years of experience with the MSPAP, standards were adopted for these assessments in July 1993. Table 6 displays the standards used in determining the progress of Maryland schools, based on each of the student performance indicators.

Table 6. Established Performance Standards for Maryland Schools

Student Performance Indicator	Standard for Receiving Satisfactory Status	Standard for Receiving Excellence Status
MFTP Assessments*		
Reading, Grade 9	95%	97%
Reading, Grade 11	97%	99%
Mathematics, Grade 9	80%	90%
Mathematics, Grade 11	97%	99%
Writing, Grade 9	90%	96%
Writing, Grade 11	97%	99%
Citizenship, Grade 9	85%	92%
Citizenship, Grade 11	97%	99%
All Tests, Grade 11	90%	96%
Yearly Attendance Rate (Grades 1-6 and 7-12)	94%	96%
Yearly Dropout Rate (Grades 9-12)	3%	1.25%
MSPAP Assessments** Grades 3, 5, 8 on all tests	70%	25%

* Percentages represent the proportion of participating students who passed the assessment.

** Percentages represent the proportion of all enrolled students (including those excused from testing or absent from school) who achieve at satisfactory or excellent performance levels. A school receives a satisfactory rating if 70% of its students achieve at satisfactory or above. A school meets the excellent standard only when 70% of its students achieve at satisfactory or above and 25% or more of its students achieve at the excellent level of proficiency.

Current Issues and Future Plans

As in other states with well-developed educational accountability systems, ongoing maintenance and evaluation become increasingly critical as the system evolves. At this point, the MSPP has identified over 50 schools that have shown inadequate progress toward improving student performance measures, or have shown no progress at all. More time is needed to determine whether this identification will push local school systems and school sites to implement their improvement plans and bring about real improvement in student performance.

Much attention currently is focused on the implementation of the High School Assessment

(HSA) Program, and its emphasis on an articulated set of Core Learner Goals. With these end-of-course exams designed to replace the current MFTP graduation exams that emphasize more basic competency skills, the state will face replacing one of the cornerstones in its current accountability system. Performance standards, both for students and for schools, will need to be established for the HSA program in order to integrate the findings of these exams into the state's overall accountability system.

This challenge of establishing and integrating new standards of student performance also applies to the Independence Mastery Assessment Program (IMAP). Discussions are currently underway to determine the best means by which results of this alternate assessment can be integrated into the overall accountability and reporting system. Can these scores be integrated into the scores collected through the MSPAP program, or should they be reported separately, with separate performance standards? The inclusion of alternate assessment results into an established accountability structure is a future challenge for those in charge of the Maryland assessment program.

A Comparison of the States' Accountability Systems

As bellwethers in creating fully inclusive educational accountability systems, Kentucky and Maryland serve as examples for other states interested in educational reform. But each state has taken a slightly different route to reforming its public education, and such differences illustrate the point that a variety of strategies can be used by states seeking change. In this concluding section, we contrast areas in which the states have taken different approaches toward building educational accountability, and also identify some underlying commonalities between the two states.

Contrasts

The establishment of benchmark performance standards. Determining benchmark performance standards is a critical component in both states' accountability systems, due to established procedures for assisting and potentially sanctioning schools that fail to progress toward these standards. But the states have taken slightly different approaches to establishing these benchmarks. In Kentucky, adequate progress is defined as a minimal increase (i.e., a 1 point increase over the school's baseline every two years) in the school accountability index calculated for each school site. In Maryland, progress is determined after a careful examination of the selected performance indicators of interest, which differ depending on grade levels served. These performance indicators are used in combination with pre-established performance standards (anticipated to be met by all local systems by the year 2000) and prior performance levels to determine a "change index" that determines whether a school has shown improvement between testing periods.

The use of test performance in graduation requirements. Another difference between these two states is their approach to using test performance for determining graduation eligibility. While Maryland's Functional Testing Program attaches high stakes (high school graduation) to a student's performance, the accountability system of Kentucky does not implement a testing program for this purpose. (Individual districts in Kentucky may have such testing requirements, but no statewide graduation testing program is mandated at the state level.)

Dissemination and reporting strategies. The state education agency in Maryland produces an annual report on the performance of each of the state's 24 local school systems, along with aggregated results for the state at large. This report offers several pieces of information, including: (1) how well each local school system performed on each selected indicator, (2) the benchmark levels of performance considered "satisfactory" or "excellent," (3) three years of information on previous district performance, and (4) two years of supporting information to help explain district performance, such as poverty indicators and proportions of students with special needs. The same information on all schools in the state is reported and published by local school systems, and is disseminated to all parents and the public. These annual reports are broadly disseminated to various audiences. Each school's level of progress can also be readily accessed in Maryland.

In Kentucky, public accountability reports contain far less detailed information on district demographic information and individual indicators of performance, and focus more exclusively on school accountability indices. The level of reporting is at the individual school level, and each individual school's level of progress can be readily accessed.

Selection of standards for alternate assessment programs. Both states are pioneers in developing alternate assessments for those students with severe disabilities who could not participate in the regular assessment programs. But interesting differences can be found in how each state dealt with the question of aligning such an alternate assessment with existing frameworks of academic standards or expectations. In the case of Kentucky, a smaller subset of 28 standards was selected from the much broader taxonomy of standards established for all students in public education. This subgrouping was seen as being applicable to students with even the most severe disabilities. In contrast, Maryland is using a broadly representative advisory group to develop and define a *separate* set of academic expectations for students eligible for participating in their alternate assessment program.

Commonalities

Several common themes emerge from this review of the scoring approaches used in the accountability systems of Maryland and Kentucky. Such themes are important because they highlight some of the approaches that other states might take to address the challenges of including all students in state accountability systems.

A premise that all students are included. The systems in both Maryland and Kentucky reflect a premise that all students count and that accountability must encompass all students. Thus, scores are aggregated and records are kept on students whose scores are not among those aggregated. These states know who is in and who is out of their aggregation and reporting systems, something not true of many other states (Erickson, Thurlow, & Thor, 1995). They actively pursue increasing the inclusion of students with disabilities in aggregation and reporting.

Accommodations are accepted, available, and not treated separately. The attitude toward accommodations in general for both Maryland and Kentucky is that they are an appropriate form of support for students with disabilities. They are not viewed as providing an unfair advantage and therefore compromising the assessment. Thus, they are not a distinguishing factor in aggregation and reporting.

"Zero" scores are used to back the belief that all students count. Both states have implemented scoring procedures that gain the attention of local decision makers, and encourages them to consider whether the decision they are making really is appropriate. This is done through two techniques in the two states, both of which essentially assign zero scores when students are kept out of the assessment. In Maryland, students who are excused or absent from school are counted in the base number of students on which scores are calculated. When a student does not take the assessment, the score entered in the system is essentially a zero. In Kentucky, students who are exempt from both the regular and the alternate assessment (like Maryland's excused students) are assigned the "novice" level as their score. This is the lowest level possible, thus, essentially a zero in that system.

Auditing procedures. In both Maryland and Kentucky, audits are built into the system, with a part of that auditing focused on students with disabilities. In Maryland, too many exemptions (the equivalent of too many students not taking the test) triggers an audit. In Kentucky, more than 2% of the student population being designated for the Alternate Portfolio system also is reason for an audit. Audit procedures convey the message that most students should be in the regular assessment system, with or without accommodations.

Reporting on all students. Both states seek to provide public information on the testing participation or performance of all their students. In Maryland, student participation is reflected in either the scores reported or in the rates of exempted, excused, and absent students. Next steps in developing performance indices for the latter two groups are under development; the performance of exempted students is to be measured by IMAP. In Kentucky, all students are in the reports that are now provided. Scores of students with disabilities are not differentiated from the scores of other students in any way. All are included in the same way in the accountability index derived for each school.

Summary

The process by which measures of student performance are used to evaluate the success of schools is never a simple or straightforward process. Efforts to measure and report progress must withstand significant psychometric and political challenges in order to succeed. The activities and decisions outlined within this report reveal the incredible efforts undertaken by these two states in their quest for using student performance data to build and maintain valid systems of statewide educational accountability. Though different in their approaches to the problem, both states have emerged as exemplars in the pursuit of establishing accountability systems that view the success of all students as critical, and the failure of any student as unacceptable.

References

Atash, N. (1994). *Establishing proficiency levels and descriptions for the 1993 Maryland school performance assessment program (MSPAP).* Rockville, MD: Westat, Inc.

Erickson, R., Thurlow, M., & Thor, K. (1995). *1994 State special education outcomes.* Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Kleinert, H. L., Kearns, J. F., & Kennedy, S. (1996). *Accountability for all students: Kentucky's Alternate Portfolio Assessment for students with moderate and severe cognitive disabilities.* Lexington, KY: University of Kentucky.

Trimble, C. S. (1994). Ensuring educational accountability. In T. R. Guskey (Ed.), *High stakes performance assessment: Perspectives on Kentucky's educational reform* (pp. 37–54). Thousand Oaks, CA: Corwin.

Trimble, C. S., & Forsaith, A. C. (1995). Achieving equity and excellence. *University of Michigan Journal of Law Reform, 28* (3), 599–653.

Ysseldyke, J., Thurlow, M., Erickson, R., Gabrys, R., Haigh, J., Trimble, S., & Gong, B. (1996). *A comparison of state assessment systems in Kentucky and Maryland with a focus on the participation of students with disabilities* (Maryland-Kentucky Report 1). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Appendix

Sample of state summary and disaggregated data, and summary data for each school system in Maryland

MARYLAND

Maryland, with a population of 4,983,900, ranks forty-second in size and nineteenth in population among the fifty states. The State Department of Education is housed in Baltimore. There are twenty-four local school systems and 1,301 public schools and centers.

STUDENT PERFORMANCE
School Year 1995 -1996

ASSESSED STUDENT KNOWLEDGE

MARYLAND FUNCTIONAL TESTS Grade 9 Status	Standard Percent Ex	Standard Percent Sat	1993** Percent Passing	1995 Percent Passing	1996 Number Taking	1996 Number Absent	1996 Number Exempt	1996 Percent Passing	1996 Ex	1996 Sat	1996 Not Met
Reading	97	85	97.4	97.1	59,765	1,072	922	97.2	✓		
Mathematics	90	80	79.2	81.1	60,012	851	884	83.0		✓	
Writing	96	90	93.5	88.3	57,288	2,957	1,254	82.5			✓
Citizenship	92	85	83.8	84.1	54,364	2,413	766	83.1			✓

MARYLAND FUNCTIONAL TESTS Grade 11 Status	Standard Percent Ex	Standard Percent Sat	1991** Percent Passing	1995 Percent Passing	1996 Number Refused	1996 Number Exempt	1996 Percent Passing	1996 Ex	1996 Sat	1996 Not Met
Reading	99	97	99.4	99.5	*	545	99.6	✓		
Mathematics	99	97	96.5	96.4	0	533	95.9			✓
Writing	99	97	97.7	98.5	0	741	97.7		✓	
Citizenship	99	97	96.4	96.1	*	495	95.5			✓
Passed All Tests	96	90	93.1	93.1	0	452	91.8		✓	

STUDENT PARTICIPATION

ATTENDANCE RATE (Yearly)	Standard Percent Ex	Standard Percent Sat	1990** Percent	1995 Percent	1996 Percent	1996 Ex	1996 Sat	1996 Not Met
Grades 1 - 6	98	94	94.2	95.0	95.1		✓	
Grades 7 - 12	98	94	90.1	91.0	91.4			✓

Students Absent	1995 Percent	1996 Percent
Fewer than 5 days	31.5	33.5
More than 20 days	14.7	13.6

DROPOUT RATE (Yearly)	Standard Percent Ex	Standard Percent Sat	1990** Percent	1995 Percent	1996 Percent	1996 Ex	1996 Sat	1996 Not Met
Grades 9 - 12	1.25	3.00	6.5	4.95	4.58			✓

MARYLAND SCHOOL PERFORMANCE ASSESSMENT PROGRAM		Standard Percent Ex	Standard Percent Sat	1993** Percent at Ex	1993** Percent at Sat	1995 Percent at Ex	1995 Percent at Sat	1996 Number Tested	1996 Number Absent/ Excused	1996 Number ESL Exemptions	1996 Number Special Education Exemptions	1996 Number Second Semester Transfers	1996 Percent at Ex	1996 Percent at Sat	1996 Ex	1996 Sat	1996 Not Met
3	Reading	25	70	---	---	4.6	34.0	58,256	2,274	832	4,013	534	4.3	35.3			✓
3	Writing	25	70	9.2	35.1	13.8	39.3	59,872	1,475	802	1,201	559	11.8	40.9			✓
3	Language Usage	25	70	9.0	29.4	10.6	43.0	58,465	824	817	3,263	540	13.4	45.2			✓
3	Mathematics	25	70	2.1	28.6	6.4	42.0	58,670	2,001	808	1,875	555	6.0	38.7			✓
3	Science	25	70	2.3	31.1	7.2	41.1	58,573	2,774	802	1,201	559	5.3	36.0			✓
3	Social Studies	25	70	1.1	31.9	2.4	38.0	60,017	1,330	802	1,201	559	2.0	29.1			✓
5	Reading	25	70	2.3	24.7	3.6	29.5	54,834	2,099	475	4,549	438	3.7	33.7			✓
5	Writing	25	70	11.7	36.8	14.8	36.7	58,950	1,203	442	1,348	452	19.1	42.3			✓
5	Language Usage	25	70	10.4	26.8	17.6	39.6	57,143	753	470	3,588	441	22.4	45.3			✓
5	Mathematics	25	70	5.8	39.5	8.8	44.7	58,146	2,007	442	1,348	452	9.6	47.8			✓
5	Science	25	70	4.0	33.3	6.9	41.2	58,179	1,974	442	1,348	452	8.1	44.8			✓
5	Social Studies	25	70	3.0	31.3	7.5	38.4	58,508	1,645	442	1,348	452	10.0	42.8			✓
8	Reading	25	70	1.2	24.6	1.5	27.6	52,062	3,713	459	1,993	531	1.6	28.6			✓
8	Writing	25	70	9.0	36.3	10.6	42.1	53,841	3,228	445	706	538	15.2	43.0			✓
8	Language Usage	25	70	8.6	36.9	17.1	52.2	54,170	2,199	448	1,405	536	16.5	52.9			✓
8	Mathematics	25	70	4.1	35.8	6.3	42.3	52,384	4,685	445	706	538	8.1	43.3			✓
8	Science	25	70	---	---	11.5	46.1	54,834	2,235	445	706	538	12.2	46.8			✓
8	Social Studies	25	70	2.8	25.9	3.8	35.9	54,358	2,711	445	706	538	4.5	36.2			✓

*Fewer Than 5 Students **Indicates Baseline Year Data KEY: Ex = Excellent, Sat = Satisfactory

SUPPORTING INFORMATION

School Year 1995-1996

Maryland

ENROLLMENT	1994-95	1995-96
Pre-Kindergarten	18,834	19,092
Kindergarten	62,341	63,232
Grades 1 - 6	375,752	381,463
Grades 7 - 12	321,124	328,655
Ungraded Special Education	12,887	13,102
TOTAL ENROLLMENT	790,938	805,544

STUDENT MOBILITY	1995		1996	
	Number	Percent	Number	Percent
Entrants	88,969	11.6	92,980	11.9
Withdrawals	97,011	12.7	97,255	12.4

STUDENTS RECEIVING SPECIAL SERVICES	1995		1996	
	Number	Percent	Number	Percent
Limited English Proficient	14,305	1.8	15,104	1.9
Title 1	53,583	6.8	62,125	7.7
Free/Reduced Price Meals	240,623	30.4	249,611	31.0
Special Education	92,175	11.7	96,543	12.0

OTHER FACTORS	1995	1996
Wealth per pupil	234,091	$232,924
Per pupil expenditure	$6,106	$6,337
Instructional staff per 1,000 pupils	57.9	57.7
Professional support staff per 1,000 pupils	8.8	8.6
Instructional assistants per 1,000 pupils	9.9	9.9
Average length of school day for pupils	6.5	6.4
Length of school year for pupils	180	180

FIRST GRADERS WITH KINDERGARTEN EXPERIENCE	1995		1996	
	Number	Percent	Number	Percent
	64,927	97.8	66,753	97.6

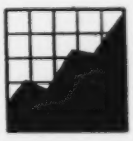
HIGH SCHOOL PROGRAM COMPLETION	1995	1996	
	Percent Completed	Number Completed	Percent Completed
University of Maryland System Requirements	49.7	21,626	50.8
Approved Occupational Program Requirements	13.1	5,412	12.7
University and Occupational Requirements	3.7	2,127	5.0

GRADE 12 DOCUMENTED DECISIONS TO:	1995	1996	
	Percent	Number	Percent
Attend a four year college	42.7	16,602	42.9
Attend a two year college	18.2	6,624	17.1
Attend a specialized school or specialized training	2.5	956	2.5
Enter employment (related to high school program)	4.2	1,556	4.0
Enter employment (unrelated to high school program)	8.3	3,282	8.5
Enter the military	3.6	1,463	3.8
Enter full-time employment and school	3.3	1,443	3.7
Enter part-time employment and/or school	12.0	4,590	11.9
Other and no response	5.3	2,212	5.7

School Improvement Notes

Maryland's Student Performance results are aggregates of the twenty-four local school systems' results. Overall, satisfactory standards were met for: Maryland Functional Tests--grade 9 in mathematics, and grade 11 in writing and passed all tests; and student attendance in grades 1 - 6. Excellent standards were met for: Maryland Functional Tests--grades 9 and 11 in reading. As a state, we did not meet the satisfactory standards for: writing and citizenship in grade 9, mathematics and citizenship in grade 11, student attendance in grades 7 - 12, dropouts, and all grades and subject areas in the Maryland School Performance Assessment Program (MSPAP). In the MSPAP, our scores decreased in grade 3 mathematics, science, and social studies; however, our scores increased in the remaining fifteen of the eighteen areas assessed.

We shall continue to work with local school systems to improve the quality of instruction to move us closer to our goal of achieving all standards.



NATIONAL CENTER ON EDUCATIONAL OUTCOMES

University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612/626-1530 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The College of Education & Human Development

UNIVERSITY OF MINNESOTA